# Mitä Gödelin epätäydellisyysteoreemoista voidaan päätellä filosofiassa

Pantsar, Markus

2011

Pantsar , M 2011 , ' Mitä Gödelin epätäydellisyysteoreemoista voidaan päätellä filosofiassa '
, Ajatus , Vuosikerta. 68 , Sivut 175-200 .

http://hdl.handle.net/10138/325200

acceptedVersion

Mitä Gödelin epätäydellisyysteoreemoista voidaan päätellä filosofiassa?

MARKUS PANTSAR
Helsingin yliopisto

## 1. Johdanto

Kurt Gödelin (1931) epätäydellisyysteoreemat ovat kenties 1900-luvun matemaattisen logiikan kuuluisin tulos. Niissä Gödel osoitti, että kaikissa ristiriidattomissa aritmetiikan (luonnollisten lukujen teorian) sisältävissä formaaleissa systeemeissä voidaan ilmaista lauseita, joita ei voida sen enempää todistaa kuin osoittaa epätodeksikaan kyseisen systeemin sisäisesti. Tällaiset ratkeamattomat *Gödel-lauseet* ovat kuitenkin *tosia* helposti hyväksyttävien totuusteoreettisten oletusten nojalla. Näin ollen Gödel-lauseiden voidaan sanoa olevan tosia lauseita, joita ei voida todistaa. Tämä ero semanttisen totuuden ja formaalin todistettavuuden välillä on inspiroinut filosofiassa monenlaisia argumentteja.

Epätäydellisyysteoreemojen sovellusten määrä ja kirjo on hämmästyttävä. Mikään muu matemaattisen logiikan tulos ei pääse lähellekään epätäydellisyyden vaikutusta, etenkin kun tarkastellaan muuta filosofiaa kuin logiikan tutkimusta. Muun muassa seuraavia asioita on argumentoitu seuraavan Gödelin epätäydellisyysteoreemoista:

(1) Kaikki ristiriidattomat aritmetiikan sisältävät formaalit matemaattiset systeemit ovat epätäydellisiä. (Gödel 1931)

(2) David Hilbertin ohjelma pyrkiä esittämään kaikki matematiikka finitistisesti on tuomittu epäonnistumaan. (esim. Smorynski 1977)

(3) On olemassa tosia matematiikan lauseita, joita ei voida todistaa, joten totuus ja todistettavuus eivät voi olla matematiikassa sama käsite. (Shapiro 1998, Ketland 1999)

(4) Matematiikka ei voi olla pelkästään ihmisten luomaa. (Gödel 1951)

(5) Joko on olemassa absoluuttisesti todistamattomissa olevia matematiikan lauseita tai ihmismieli ylittää kaikkien koneiden todistuskyvyn. (Gödel 1951)

(6) Mikään mekaaninen malli ei voi kuvata täydellisesti ihmismieltä. (Lucas 1961)

(7) Vahva tekoäly ei ole mahdollista, eli tietokone ei voi koskaan ajatella kuin ihminen. (Penrose 1989, 1994)

(8) Mikään matemaattinen todistus ei ole enää tarkalleen ottaen loogisesti pätevä, koska emme voi osoittaa ristiriidattomuutta käytetylle todistussysteemille. (Nagel & Newman 1959)

(9) Kaiken selittävät teoriat eivät ole mahdollisia fysiikassa. (Dyson)

(10) Epätäydellisyys antaa meille syitä uskoa Jumalaan. (Graves, Guillen)

(11) Moninaiset sovellukset mm. sosiologiassa (Debray 1980) ja runoudessa (Kristeva 1969).

Tässä listassa ainoastaan kohdat (1) ja (2) ovat kiistattomia. Näistä ensimmäinen on Gödelin ensimmäisen epätäydellisyysteoreeman sisältö ja toinen epätäydellisyysteoreemojen välitön matemaattinen seuraus. Kaikki muut kohdat ovat enemmän tai vähemmän kiistanalaisia. Tarkoituksenani on käydä nämä Gödel-tulkinnat järjestelmällisesti läpi ja siten selventää millaisia sovelluksia epätäydellisyydestä voidaan filosofiassa tehdä. Edellä esitetty johtopäätösten joukko ei ole missään nimessä kaikenkattava: gödeliläiset päätelmät ja virhepäätelmät ovat tuhansia ja taas tuhansia sivuja tuottanut filosofian ala. Mutta uskon, että tässä esitellyillä johtopäätöksillä saamme tarpeeksi laajan kuvan siitä, mitä Gödelin epätäydellisyysteoreemoista voidaan päätellä filosofiassa – ja mikä kenties vielä tärkeämpää, mitä *ei* voida.

## 2. Gödelin epätäydellisyysteoreemat

Vaikka emme mene tässä syvemmälle Gödelin epätäydellisyysteoreemojen todistukseen, on syytä tarkastella lyhyesti niiden sisältöä ja niitä ehtoja, joiden nojalla epätäydellisyys pätee. Varsinkin siinä vaiheessa kun tarkastellaan radikaalimpia epätäydellisyyden sovelluksia, on erittäin tärkeää muistaa, että Gödelin todistus vaatii selkeät alkuehdot ollakseen pätevä. Epämuodollisesti Gödelin ensimmäisen epätäydellisyysteoreeman sisältö on seuraava. Olkoon $\mathbf{S}$ jokin mielivaltainen formaalin matematiikan systeemi, joka täyttää seuraavat ehdot:

1) $\mathbf{S}$ on ristiriidaton

2) $\mathbf{S}$ sisältää aritmetiikan eli luonnollisten lukujen (0, 1, 2, …) teorian.

Tällöin systeemin $\mathbf{S}$ kielessä on olemassa lause $\varphi$, jolle pätee seuraava:

i) $\mathbf{S} \nvdash \varphi$

ii) $\mathbf{S} \nvdash \neg\varphi$

Eli emme voi johtaa systeemissä **S** lausetta $\varphi$, mutta emme myöskään sen negaatiota. Olkoon **S** kuten edellä. Gödelin toinen epätäydellisyysteoreema osoittaa, että lisäksi:

iii) $\mathbf{S} \nvdash \mathrm{Con}_{\mathbf{S}}$,

jossa $\mathrm{Con}_{\mathbf{S}}$ on lause joka ilmaisee **S**:n ristiriidattomuuden.

Gödelin todistuksessa lause $\varphi$ on muotoiltu ns. diagonaalimenetelmällä seuraavasti: $\varphi \leftrightarrow \neg Pr_S(\ulcorner\varphi\urcorner)$, eli semanttisesti puhuen, lause $\varphi$ on tosi systeemissä **S** jos ja vain jos sitä ei voida todistaa kyseisessä systeemissä. Merkintä $\ulcorner\varphi\urcorner$ tarkoittaa lauseen $\varphi$ *Gödel-numeroa*, joka on todistuksessa tarvittava koodaus. Mutta Gödel todistaa, että lausetta $\varphi$ ei todellakaan voida todistaa systeemissä **S**. Siispä $\varphi$ on tosi, muttei todistettava. Gödel itse asiassa uskoi tämänkaltaisen semanttisen todistuksen olevan pätevä, mutta käytti syntaktista todistusta David Hilbertin formalistisen vaikutuksen takia. Tulemme näkemään, että semanttinen (eli *tarskilainen*) totuus on vieläkin kiistanalainen käsite, mutta syntaktisen todistuksen ansiosta Gödel osoitti formaalien systeemien epätäydellisyyden kiistattomasti. Koska **S** oli mielivaltainen, tiedämme että kaikki aritmetiikan sisältävät ristiriidattomat formaalit systeemit ovat epätäydellisiä. Lisäksi ne eivät voi todistaa omaa ristiriidattomuuttaan.

Tämä esitys on epäformaali ja äärimmäisen tiivistetty, mutta siitä selviää silti kolme filosofiassa tärkeää seikkaa Gödelin epätäydellisyysteoreemoista. Ensimmäiseksi, teoreemojen varsinainen matemaattinen sisältö. Hilbertin vaikutuksesta matematiikassa oli pyrkimys näyttää, että kaikki matematiikka voidaan palauttaa finitistisiin aksioomiin ja päättelysääntöihin, lopulta kenties yhteen formaaliin matemaattiseen systeemiin. Gödel romutti nämä toiveet näyttämällä, että tällaiset ristiriidattomat systeemit ovat epätäydellisiä heti kun ne sisältävät lukuteorian, jota yleisesti pidetään matematiikan yhtenä perustavimmista osa-alueista. Tämä on tietenkin jo itsessään merkittävä tulos matematiikanfilosofiassa.

Toiseksi, meillä on hyvin yksinkertainen syy uskoa, että tällainen ratkeamaton Gödel-lause on itse asiassa *tosi*. Filosofiassa tarvitsemme totuusteorian tämän tuloksen tueksi, mutta on ymmärrettävä intuitio ajatella Gödel-lauseita todistamattomina, mutta tosina.

Kolmanneksi, saamme ehdot, joilla epätäydellisyys pätee. Tulemme huomaamaan, että näitä ehtoja pidetään joskus filosofiassa vähemmän tärkeinä. Mutta todellisuudessa ne ovat kaikki kaikessa, kun alamme soveltaa Gödelin tuloksia filosofiassa. Epätäydellisyysteoreemat pätevät *ainoastaan* ristiriidattomissa aritmetiikan sisältävissä formaaleissa systeemeissä.

**3. Semanttinen argumentti**

Edellä nähtiin, että meillä on hyvä syy ajatella todistamattomissa olevien Gödel-lauseiden olevan tosia. Mutta tämä ei sellaisenaan merkitse mitään ennen kuin olemme tarkentaneet, mitä totuudella tarkoitetaan. Selvästikään totuudella ei tässä tarkoiteta todistettavuutta kyseisessä formaalissa systeemissä. Systeemin **S** Gödel-lause (merkitään G(**S**)) on osoitettu todistamattomaksi, joten ollakseen tosi, totuuden täytyy tarkoittaa jotain muuta kuin pelkkää todistettavuutta.

Vaikka lauseen todistettavuus ei siis seuraa sen totuudesta, varmasti haluamme matemaattisten systeemiemme olevan sellaisia, että lauseen totuus seuraa sen todistettavuudesta, mitä ikinä tarkoitammekaan totuudella. Jos formaalin systeemin aksioomat ovat tosia ja sen päättelysäännöt säilyttävät totuuden, ovat kaikki systeemissä todistettavat lauseet tosia. Tätä ehtoa kutsutaan *eheydeksi*. Nyt meitä kiinnostava kysymys kuuluu: mitä tapahtuu kun formaali systeemi on ehyt, mutta epätäydellinen, eli kun kaikki siinä todistettava on totta, mutta siinä ei voida todistaa kaikkia systeemissä ilmaistavia totuuksia?

Tätä tarkoitusta varten tarvitsemme *totuuspredikaatin*, eli joudumme laajentamaan formaalia systeemiämme. Alfred Tarski osoitti vuonna 1933 (ks. Tarski 1935), ettei mikään ilmaisuvoimaltaan kyllin rikas formaali klassisen logiikan kieli voi sisältää omaa totuuspredikaattiaan. Siksi emme voi laajentaa systeemiä **S** totuuspredikaatilla siten, että saatu systeemi **S’** sisältäisi oman totuuspredikaattinsa. Tämä synnyttää valehtelijan paradoksin: muotoa ”tämä lause on epätosi” olevan lauseen totuusarvoa ei voida ratkaista, joten mikäli haluamme totuuspredikaatin olevan täydellinen, emme voi sallia kielen puhuvan sen omien lauseidensa totuudesta tai epätotuudesta. Sen sijaan laajennettu systeemi **S’** voi sisältää totuuspredikaatin *alkuperäiselle* systeemille **S**.

Mitä voimme odottaa tällaiselta totuuspredikaatilta? *Deflationistien* (kuten Paul Horwich (1998) ja Hartry Field (1999)) mukaan totuus on tyhjä käsite, jonka ainoa tehtävä on toimia Tarskin T-lauseissa: ”p” on totta jos ja vain jos p. Koska deflationismissa totuus ajatellaan tyhjänä käsitteenä, ei totuuspredikaatin lisäämisestä saa luonnollisestikaan seurata *uusia* totuuksia. Toisin sanoen, totuuspredikaatin tulee olla *konservatiivinen*. Kun lisäämme pelkästään totuuspredikaatin formaaliin systeemiin **S**, deflationistisen totuusteorian mukaan meidän ei pitäisi pystyä uudessa systeemissä **S’** todistamaan yhtään sellaista systeemin **S** lausetta, jota emme voi todistaa **S**:ssä itsessään.

Lisätään nyt totuuspredikaatti ristiriidattomaan aritmetiikan sisältävään formaaliin systeemiin **S**. Jos **S** on ehyt, on jokainen **S**:ssä todistettava lause tosi **S’**:ssa. Siis systeemissä **S’** pätisi seuraava eheysominaisuus:

$$\text{(E)} \qquad \mathbf{S}' \vdash \forall x(\mathrm{Pr}(x) \to Tr(x)),$$

missä *Pr* on systeemin **S** todistettavuuspredikaatti. Otetaan nyt jokin mielivaltainen ristiriitainen aritmetiikan lause, vaikkapa $\ulcorner 0 = 1 \urcorner$ (tarvitsemme Gödel-numerointia, sillä systeemissä **S’** lause 0 = 1 voi olla muotoiltu eri tavalla). Koska **S** on oletettu ristiriidattomaksi, tällöin jonkin systeemin **S’** T-lauseen täytyy kertoa meille $\mathbf{S}' \vdash \neg Tr\ulcorner 0 = 1 \urcorner$. Kontraposition lain mukaan tällöin $\mathbf{S}' \vdash \neg Pr\ulcorner 0 = 1 \urcorner$. Mutta $\neg Pr\ulcorner 0 = 1 \urcorner$ on ekvivalentti lauseen $\mathrm{Con}_{\mathbf{S}}$ kanssa. Jos totuuspredikaatti olisi konservatiivinen, pätisi $\mathbf{S} \vdash \mathrm{Con}_{\mathbf{S}}$, sillä $\mathrm{Con}_{\mathbf{S}}$ on systeemin **S** lause. Mutta Gödelin toisen epätäydellisyysteoreeman mukaan $\mathbf{S} \nvdash \mathrm{Con}_{\mathbf{S}}$, joten saavumme ristiriitaan.

Tässä päättelyssä meillä oli kaksi oletusta, joista vähintään toinen täytyy nyt hylätä. Koska lause (E) kuulostaa ehdottomasti siltä, mitä toivomme adekvaatilta totuuspredikaatilta, jää ainoaksi vaihtoehdoksi hylätä oletus totuuspredikaatin konservatiivisuudesta. Siispä lisäämällä pelkän totuuspredikaatin formaaliin systeemiin **S** voimme johtaa sellaisia **S**:n totuuksia, jotka eivät ole todistettavissa **S**:ssä. Näin ollen deflationistinen totuusteoria näyttäisi kaatuvan: yksinkertainen tarskilainen totuus on *substantiaalinen*, se antaa meille *uusia* totuuksia. Tämä on ns. *semanttinen argumentti* deflationismia vastaan, ja sen ovat esittäneet oleellisesti samassa muodossa Stewart Shapiro (1998) ja Jeffrey Ketland (1999).

Onko semanttinen argumentti pätevä epätäydellisyysteoreemojen sovellus filosofiassa? Asiasta käydään yhä paljon keskustelua. Ensimmäinen asia mikä tulee selventää, on laajennuksessa käytettävä totuuspredikaatti. Jotta pystymme johtamaan (E):n kaltaisen yleistyksen, täytyy totuuspredikaatin sisältää matemaattinen induktio predikaatin sisältävien lauseiden suhteen. Field (1999) esitti kritiikin, jonka mukaan matemaattinen induktio ei ole *totuuden*, vaan *aritmetiikan* ominaisuus, joten sitä ei tule sisällyttää totuuspredikaattiin. Tämä ei kuitenkaan ole helposti perusteltavissa, sillä deflationistin mukaan *mikään* ei ole pelkästään totuuden ominaisuus. T-lauseet riippuvat aina niiden sisällöstä ja on luonnollista ajatella, että *luonnollisia lukuja* koskevat T-lauseet tottelevat matemaattista induktiota, joka on luonnollisten lukujen keskeinen - jopa ne määrittelevä, kuten Bertrand Russellilla (1920) - ominaisuus.

Tarskilainen totuuspredikaatti ei siis ole konservatiivinen, mutta pitääkö sen edes olla sitä? Volker Halbach (2001) on argumentoinut, että deflationistisen totuuden ei tarvitse olla konservatiivinen – riittää, että se luettelee kaikki **S**:n teoreemojen T-lauseet. Mahdolliset muut totuudet ovat vain ”toissijainen vaikutus”. Tämä ei kuitenkaan ole ongelmatonta, sillä millä perusteella ensisijaiset ja toissijaiset totuudet voidaan erotella, jos ne seuraavat samasta totuuspredikaatista? Lisäksi Gödel-lauseet *ovat* selvästikin tosia systeemissä **S’**, joten rajatessamme totuuspredikaatin vaikutusalueeksi **S**:n teoreemat, luovumme tärkeästä tuloksesta. Samankaltaisella argumentilla Jody Azzouni (1999) pyrki rajoittamaan deflationistisen totuuden koskemaan **S**:n teoreemoja sillä perusteella, että deflationisti ei tarvitse lauseen (E) kaltaisia yleistyksiä. En pidä näitä vasta-argumentteja kovinkaan hedelmällisinä. Totuuspredikaatin soveltuvuusalueen rajoittaminen tuntuu epäilyttävän mielivaltaiselta: tulee muistaa, että Gödel-lauseiden totuus seuraa pelkästään formaalista systeemistä **S** ja tarskilaisesta totuuspredikaatista. On vaikeaa nähdä mitään ongelmaa kummassakaan.

David Isaacson (1987) on esittänyt lupaavamman erottelun aritmeettisen teorian kuten ensimmäisen kertaluvun Peano aritmetiikan (**PA**) teoreemojen ja muiden **PA**:n totuuksien välille. Hänen mukaansa ainoastaan **PA**:n teoreemat ovat *aritmeettisia* totuuksia, sillä niiden ymmärtämisessä ei mennä aritmetiikan käsitteitä syvemmälle. On syytä muistaa, että luonnolliset luvut toimivat kahdessa roolissa Gödelin epätäydellisyysteoreemoissa. Toki olemme kiinnostuneita luonnollisista luvuista niiden tavanomaisessa mielessä, mutta myös *Gödel-numeroina*. Koska lauseen G(**PA**) ratkeamattomuuden todistamisessa ja sen totuuden johtamisessa käytämme Gödel-numerointia, olemme Isaacsonin mukaan käyttäneet ”piilotettua korkeamman kertaluvun käsitettä”.

Tämä on erittäin aiheellinen huomio, sillä epätäydellisyysteoreemojen todistuksessa todellakin käytämme hyväksemme tietoa, että jokainen formaalin systeemin lause voidaan koodata luonnollisiksi luvuiksi. Saamme koodauksella paljon lisää ilmaisuvoimaa, mutta selvästikään emme käytä siinä luonnollisia lukuja samalla tavalla kuin aritmetiikassa. Siksi Isaacson erottelee aritmeettiset totuudet (eli aritmetiikan teoreemat) korkeamman kertaluvun totuuksista, joihin G(**PA**) kuuluu. Näin ollen aritmeettisten totuuksien joukko voi olla täydellinen, vaikka **PA**:ssa ilmaistavien totuuksien joukko ei sitä olekaan. Leon Horsten (2001) on käyttänyt samankaltaista argumenttia joukko-opin totuuksista.

Isaacsonin argumentti on mielenkiintoinen, mutta onko sillä voimaa deflationismin puolustuksena?[1] Vaikka hyväksymme, että deflationistinen totuusteoria voi sisältää kaikki aritmetiikan totuudet, säilyy sama ongelma korkeamman kertaluvun totuuksilla, joilla deflationistinen totuusteoria ei ole konservatiivinen.

Totuuden konservatiivisuus tuntuu olevan deflationismin kompastuskivi, mutta on myös muita tapoja kiistää semanttisen argumentin pätevyys. Neil Tennant (2002) on (Solomon Fefermania (1991) seuraten) näyttänyt, että voimme johtaa Gödel-lauseen myös ilman totuuspredikaattia, laajentamalla systeemiä **S** sopivalla eheysperiaatteella $Pr_S([x]) \to x$. Tällainen laajennettu systeemi **S’** pystyy sisältämään teoreemana systeemin **S** Gödel-lauseen, eikä näin ollen totuuspredikaattia tarvita. Tennantin argumentti on pätevä, joten kysymykseksi tulee eri laajennusten oikeutus: onko meillä enemmän perusteita laajentaa formaalia systeemiä totuuspredikaatilla vai suoralla eheyslauseella?

Omassa tutkimuksessani (Pantsar 2009) olen pyrkinyt siirtämään painopisteen nimenomaan tähän ongelmaan. Kuten näimme, eheysperiaate seuraa luonnollisesti adekvaatista totuuspredikaatista, joten sen hyväksyminen on kiistatonta. Mutta on syytä huomata ero sen välillä, että eheys *seuraa* totuuden käsitteestä ja eheysperiaate *oletetaan* suoraan. Yksinkertaistettuna kysymys on siitä, olemmeko valmiita tekemään laajennuksen mihin formaaliin systeemiin tahansa. Jos **S** on jokin mielivaltainen formaali systeemi, haluammeko laajentaa sitä tarskilaisella totuudella? Tämä riippuu tietenkin systeemin aksioomista ja päättelysäännöistä. Kun suoritamme tarskilaisen laajennuksen, väitämme, että

[1] On syytä huomata, että Isaacson itse ei argumentoinut deflationismin puolesta.

**S**:n kaikki teoreemat ovat tosia. Tietenkään emme halua tehdä tätä *kaikille* formaaleille systeemeille, vaan pelkästään niille joiden uskomme olevan tosia, mitä ikinä tarkoitammekaan totuudella. Sen sijaan on epäselvää, mitä tarkoitamme kun laajennamme formaalia systeemiä eheysperiaatteella. Emme varmastikaan halua väittää kaikkien formaalien systeemien olevan eheitä, mutta millä perusteella voimme erotella systeemien välillä?

Joko ajaudumme formaalien systeemien mielivaltaisuuteen tai tarvitsemme jonkinlaisia kriteerejä siitä, millaisia matematiikan teoreemoja haluamme väittää. Tällaiset ”väitettävyyskriteerit” ovat kuitenkin huomattavan paljon epäintuitiivisempia kuin totuuden käsite. Pyrin väitöskirjassani osoittamaan, että kun tarkastelemme matematiikkaa laajempana ilmiönä kuin pelkästään formaaleina systeemeinä, huomaamme totuuspredikaatin olevan erittäin luonnollinen laajennus formaalille matematiikalle.

Onko semanttinen argumentti totuuden substantiaalisuudesta siis pätevä? Vastaukseni on kyllä, mutta se on tärkeää asettaa oikeaan kontekstiin. Kun tarkastelemme kaikkea sitä esiformaalia matematiikkaa, joka on edeltänyt nykyisiä formaaleja systeemejä, tulee selväksi, ettei matematiikka ole mielivaltaista. Voimme ymmärtää matemaattisen totuuden filosofisesti eri tavoilla, mutta lähes kaikki matemaatikot ovat yhtä mieltä siitä, mitkä lauseet *ovat* matemaattisia totuuksia. Gödel osoitti, että ristiriidattomat aritmeettiset systeemit ovat epätäydellisiä, mutta matemaattisten totuuksien joukkoon hänen todistuksensa ei vaikuta. Se, että yksittäinen formaali systeemi ei voi saavuttaa kaikkia matematiikan totuuksia on mielenkiintoinen tulos, mutta mitään yksittäistä matemaattista totuutta se ei tee todistamattomaksi. Gödel-lauseiden totuus onkin vain oire tästä laajemmasta ilmiöstä: tarkastelemme formaaleja systeemejä aina meta-systeemeissä, lopulta epäformaalissa matematiikan kielessä. Tässä kontekstissa on täysin luonnollista, että epätäydellisyydestä seuraa tosia todistamattomia lauseita, kun tarkastelemme *yksittäisiä* formaaleja systeemejä.

Siispä totuus ja todistettavuus ovat eri käsite, joten gödeliläinen sovellus (3) on pätevä. Mutta todistamattomissa olevien matematiikan totuuksien olemassaolo ei seuraa tästä. Totuus ja todistettavuus ovat tässä tarkastelussa formaaleja käsitteitä, joilla on edellä esitetyt ominaisuudet Gödel-lauseiden suhteen. Mitään laajempaa matematiikanfilosofista johtopäätöstä tästä ei pidä tehdä.

## 4. Gödel ja platonismi

Mitä mieltä Gödel itse oli epätäydellisyysteoreemojen filosofisesta merkityksestä? On hieman kiistanalaista, mikä Gödelin matematiikanfilosofinen kanta oli vuonna 1931, mutta myöhemmällä iällään Gödel oli kiistatta platonisti ja löysi epätäydellisyysteoreemoista argumentin filosofiselle näkökannalleen matemaattisten totuuksien objektiivisuudesta. Yleensä Gödel oli varovainen julkaisemaan filosofisia näkemyksiään, mutta vuoden 1951 kuuluisassa Gibbs-luennossa Gödel esitti selkeän epätäydellisyysteoreemoihin perustuvan filosofisen argumentin. Jos pyrimme esittämään kaiken matematiikan *yhtenä* formaalina systeeminä, on siinäkin (jos se on ristiriidaton) Gödel-lause, jota ei voi todistaa tässä systeemissä, mutta jonka totuuden me näemme. Yleisesti hyväksytyn ns. Church-Turingin teesin mukaan formaali systeemi voidaan samaistaa ns. Turingin koneen[2], ja siten mekaanisten koneiden kanssa. Näin ollen formaalien systeemien epätäydellisyyden voi laajentaa koskemaan kaikkien algoritmien epätäydellisyyttä: mikään mekaaninen kone ei voi antaa meille kaikkia matematiikan totuuksia. Gödelin mukaan tästä seuraa, että:

> [Joko] ihmismieli (jopa puhtaan matematiikan alueella) ylittää äärettömästi jokaisen äärellisen koneen ratkaisuvoiman tai muuten on olemassa absoluuttisesti ratkaisemattomissa olevia [aritmetiikan] ongelmia. (Gödel 1951, s. 13)

Gödelin mukaan molemmat vaihtoehdot ovat suoraan materialistista filosofiaa vastaan. Ensimmäisessä tapauksessa ihmismieltä ei voi palauttaa aivoihin, jotka ovat äärellinen kone. Toisessa tapauksessa matematiikka ei voi olla meidän luomuksemme, koska luojan täytyy tuntea välttämättä kaikki luomuksensa ominaisuudet. Niitä ei voi olla muita kuin ne, jotka luoja on sille antanut. (Gödel 1951, ss. 15-16)

Gödelin perustelut johtopäätöksilleen ovat kenties ajan hampaan syömiä, mutta samat argumentit ovat eläneet aktiivista elämää myöhempinä vuosikymmeninä. Ensimmäistä vaihtoehtoa Gödel piti yhteensopivana parhaan aivotutkimuksen ja neurofysiologisen tiedon kanssa. Tästä tuskin moni on nykyisin samaa mieltä. Toinen vaihtoehto tuntuisi kumpuavan vahvasti sodanjälkeisten mekaanisten koneiden tasosta. Senaikaiset tietokoneet olivat yksinkertaisten algoritmien pohjalta toimivia helpommin hallittavia järjestelmiä. Nykyisten tietokoneohjelmien kompleksisuus voi usein aiheuttaa tilanteita, joissa ohjelmalla on ominaisuus, jota sen luoja ei tunne.

[2] Turingin kone on helpoiten ymmärrettävissä nykyisten digitaalisten tietokoneiden idealisaationa.

Toki on muistettava, että Gödel puhui nimenomaan matematiikasta, kun hän käytti puolustuksenaan analogiaa koneisiin (p.18). Vaikka tämä analogia on kenties vanhentunut, on olemassa mahdollisuus että formaalit matematiikan systeemit eivät ole samalla tavalla kompleksisia edes silloin kuin niiden tarkoituksena on sisältää *kaikki* matematiikka.

En kuitenkaan pidä Gödelin argumenttia kovinkaan vakuuttavana, vaikka unohtaisimme analogian koneisiin. Ensinnäkin, on mahdollista että matematiikka on ihmisen luomus, mutta se on luomus, jota ei voi ilmaista yhdellä formaalilla systeemillä. Matematiikka on jakautunut eri haaroihin, jotka ovat kehittyneet rinnakkaisesti. Voi olla, että tätä kokonaisuutta ei voida enää palauttaa yhdeksi aksiomaattiseksi systeemiksi. Voimme tuntea kaiken matematiikassa, mutta emme voi esittää sitä yhdessä formaalissa systeemissä, mikä olisi yhteensopivaa ratkeamattomien lauseiden olemassaolon kanssa. Toiseksi, formaalien systeemien epätäydellisyys seuraa niiden ilmaisuvoimasta, erityisesti siitä että niissä voidaan ilmaista tietynlainen itseensä viittaava lause. Eikö ole mahdollista, että samanlainen epätäydellisyys seuraa kaikesta matemaattisesta ajattelusta ja kielenkäytöstä? Voihan olla, että käytämme mitä tahansa tarpeeksi rikasta kieltä, siinä ei voida ilmaista kaikkia sen kielen totuuksia. Kolmanneksi, on mahdollista, että jokainen tarpeeksi rikas formaali systeemi on epätäydellinen muussakin kuin Gödelin todistamassa mielessä. Kenties kaiken matematiikan sisältävä formaali systeemi sisältäisi absoluuttisesti todistamattomia lauseita, joita ei voida nähdä todeksi. Olemme niin kaukana tällaisista systeemeistä, että tällainen argumentti on puhtaan spekulatiivinen, mutta niin on myös sen käänteisargumentti. Meillä ei ole mitään syytä olettaa, että kaiken ihmisten tunteman matematiikan sisältävä formaali systeemi olisi täydellinen, vaikka emme huomioisi Gödel-lauseita.

Epätäydellisyysteoreemat eivät toki olleet Gödelille ainoa, tai edes tärkein argumentti platonismin puolella. Tässä en mene muihin argumentteihin, vaan tyydyn toteamaan, että epätäydellisyysteoreemat eivät anna meille vakuuttavaa syytä ajatella matematiikan olevan platonistista.

## 5. Lucas ja Penrose

Gödel itse keskittyi disjunktiivisen argumenttinsa toiseen osaan eli platonismiin, mutta argumentin ensimmäinen osa on noussut pinnalle enemmän tai vähemmän tasaisin väliajoin eri muodoissa. Vuonna 1961 John Lucas pyrki osoittamaan epätäydellisyysteoreemoihin nojaten, että mikään mekaaninen malli ei voi kuvata täydellisesti ihmismieltä. Hänen argumenttinsa oli yksinkertaisesti se, että jos oletamme jonkin Turingin koneen mallintavan täydellisesti ihmismieltä, on tällä Turingin koneella (tai tarkkaan ottaen sitä vastaavalla formaalilla systeemillä) oma Gödel-lauseensa, jonka totuuden me voimme nähdä, mutta kone ei. Siispä kyseinen Turingin kone ei mallinnakaan ihmismieltä täydellisesti.

Tässä vaiheessa meidän on syytä muistaa epätäydellisyysteoreemojen todistuksessa käytetyt ehdot. Epätäydellisyysteoreemat pätevät

(1) formaaleissa systeemeissä,

jotka ovat

(2) tarpeeksi ilmaisuvoimaisia ilmaisemaan aritmetiikan

ja

(3) ristiriidattomia.

Ehto (1) on se mitä Lucasin argumentissa tutkitaan: voidaanko ihmismieli mallintaa täydellisesti formaalilla systeemillä? Ehdon (2) voidaan olettaa pätevän. Mutta miten on ehdon (3) laita? Jos oletamme, että formaali systeemi **S** voisi mallintaa ihmismielen, voimmeko olettaa, että se olisi tällöin ristiriidaton? Mikäli voimme, silloin oletamme *ihmismielen* olevan ristiriidaton. Ihmisaivot sisältävät noin sata miljardia neuronia, joista jokaisella on keskimäärin 7000 synaptista yhteyttä toisiin neuroneihin. Lucasin argumentti on toki tarkoitettu *kumoamaan* ajatus ihmismielestä formaalina systeeminä, mutta se on parhaimmillaankin pätevä ainoastaan vasta-argumenttina mekaanista mielen mallia vastaan. Siispä argumentti on relevantti vain jos ensin oletetaan, että jokin tietty formaali systeemi voisi edes *potentiaalisesti* mallintaa tämän valtavan synapsien kirjon. Tämä on jo hyvin ongelmallista, mutta sen lisäksi meidän pitäisi olettaa, että tällainen systeemi on *ristiriidaton*.

Roger Penrose (1989, 1994) on esittänyt samankaltaisen argumentin ns. *vahvaa tekoälyä* vastaan. Vahva tekoäly tarkoittaa sitä, että kone voisi periaatteessa ajatella kuin ihminen. Penrosen argumentti näyttää lupaavammalta, sillä siinä keskitytään ihmismielen ja –aivojen sijaan niiden tuotokseen, ihmis*ajatteluun*. Aivot voivat olla peruuttamattomasti liian monimutkainen elin selvitettäväksi, emmekä mahdollisesti (jopa hyvin todennäköisesti) koskaan voisi tietää mallintaako jokin formaali systeemi täydellisesti ihmisaivoja ja onko se ristiriidaton. Ajattelu on tähän verrattuna yksinkertaisempi ilmiö ja helpompi tutkia. Penrosen päätarkoituksena onkin näyttää, että matemaattinen ajattelu ei ole algoritmista, eikä siten mallinnettavissa tietokoneella. Ihminen voi nähdä Gödel-lauseiden totuuden, vaikka mikään tietokone ei siihen pystyisikään. Kuten Gödel, myös Penrose käyttää tätä argumenttina platonismin puolesta.

Penrosen argumentin voitaisiin ajatella olevan pätevä, jos voisimme rajata kaiken matemaattisen ajattelumme, formalisoida sen systeemiksi **S**, osoittaa että **S** on ristiriidaton, ja osoittaa ettemme käytä **S**:n ulkopuolisia keinoja nähdessämme G(**S**):n totuuden. Tämä mahdollisuus kaatuu kuitenkin viimeiseen kohtaan. Jos **S** on kuvailemamme formaali systeemi, meidän täytyy siirtyä systeemiin **S'** nähdäksemme G(**S**):n totuuden. Mutta nyt systeemillä **S'** on oma Gödel-lauseensa G(**S'**), jonka totuuden nähdäksemme joudumme siirtymään systeemiin **S''**. Emme pääse tästä hierarkiasta eroon muuten kuin tarkastelemalla *koko* ihmisajattelua formaalina systeeminä. Mutta koko ihmisajattelu on äärimmäisen monimutkainen ilmiö, kenties yhtä monimutkainen kuin ihmisaivojen rakenne. Ajaudumme samaan ongelmaan kuin Lucas: Gödelin epätäydellisyysteoreemat pätevät ainoastaan ristiriidattomille formaaleille systeemeille, eikä meillä ole mitään syytä ajatella, että vaikka tietäisimme jonkin formaalin systeemin mallintavan ihmisajattelua (mikä on itsessään jo utopiaa), voisimme osoittaa tämän systeemin olevan ristiriidaton.

Emmekö me kuitenkin oleta ajattelumme olevan ristiriidatonta? Ristiriitaisista lausejoukoista voidaan päätellä mitä tahansa, joten meidän on pakko olettaa ajattelumme päteväksi, mikäli haluamme päätellä mitään ei-triviaalia maailmasta. Mutta tähän tarkoitukseen riittää olettaa, että tietyt ajattelun osa-alueet ovat ristiriidattomia. Kenties ammattimatemaatikon ajattelu aritmetiikassa voi olla oleellisesti ristiriidatonta, mutta meillä ei ole mitään syytä olettaa, että hänen *kaikki* ajattelunsa olisi ristiriidatonta siinä eksaktin formaalissa mielessä, mitä Lucasin ja Penrosen argumentit vaativat. On järkevää uskoa jonkinasteiseen ihmisajattelun pätevyyteen, mutta Lucas ja Penrose vaativat valtavasti enemmän. Siksi gödeliläiset argumentit mekaanista mielen mallia ja vahvaa tekoälyä vastaan saavat toistaiseksi konditionaalisen muodon ”jos formaali systeemi **S** mallintaa ihmisajattelun ja se on ristiriidaton, niin on olemassa lause jota ei voida todistaa, mutta jonka voi nähdä todeksi”. Tässä konditionaalissa etujäsen on äärimmäisen spekulatiivinen.

## 6. Matemaattisten systeemien epäpätevyys

On huomattavan yleinen tulkinta Gödelin epätäydellisyysteoreemoista, että matemaattinen varmuus kärsi niiden seurauksena vakavan kolauksen. Ernst Nagel ja James Newman (1959) kirjoittavat paljon luetussa epätäydellisyysteoreemojen esityksessään:

> [Gödel osoitti] että on mahdotonta osoittaa sisäinen looginen ristiriidattomuus hyvin suurelle luokalle deduktiivisia systeemejä – esimerkiksi peruslukuteorialle – ellemme ota käyttöön niin kompleksisia päättelyperiaatteita, että niiden sisäinen ristiriidattomuus on yhtä avointa epäilylle kuin systeemien itsensä. (Nagel & Newman 1959, p. 6)

Mutta onko meillä syytä epäillä tärkeimpien matemaattisten systeemien ristiriidattomuutta? Ja vielä tärkeämpää aiheemme kannalta, onko meillä syytä epäillä niitä Gödelin epätäydellisyysteoreemojen pohjalta?

On tärkeää huomata, että Gödelin teoreemat eivät millään tavalla viittaa siihen, että formaalit systeemin olisivat *ristiriitaisia*. Kuten olemme huomanneet, ne eivät edes päde ristiriitaisille systeemeille. Lyhyesti sanottuna: epätäydellisyysteoreemoilla ei ole mitään tekemistä sen kanssa, pidämmekö joitain matemaattisia systeemejä ristiriidattomina vai emme. Matemaatikkojen usko aritmetiikan tai joukko-opin ristiriidattomuuteen ei ole horjunut millään tavalla Gödelin tulosten perusteella. Itse asiassa ne voidaan jopa tulkita asian kannalta positiiviseksi kehitykseksi: selittämätön ristiriidattomuustodistuksen puute Peanon aritmetiikalle ja valinta-aksioomalla vahvistetulle Zermelo-Fraenkelin joukko-opille (ZFC) olisi voinut saada aikaan epäilyjä niiden ristiriidattomuudesta. Nyt *tiedämme* että jos ne ovat ristiriidattomia, niiden sisäistä ristiriidattomuustodistusta ei voi olla olemassa.

Torkel Franzén (2005) on argumentoinut vakuuttavasti sitä vastaan, että Gödelin epätäydellisyysteoreemat olisivat tuoneet lisää epävarmuutta matematiikkaan. Hän kysyy, mikä tarkoitus jonkin formaalin systeemin sisäisellä ristiriidattomuustodistuksella olisi:

> ”Mitä hyötyä olisi ZFC:n ristiriidattomuustodistuksella, joka olisi annettu ZFC:n sisällä? Koska ZFC:n ristiriidattomuus on juuri se, mikä on kyseenalaistettu, ei ole mitään syytä ajatella tällaisella todistuksella olevan painoarvoa.” (Franzen 2005, s. 105)

Tämä on erittäin tärkeä huomio. Jos epäilemme ZFC:n olevan ristiriitainen, emme missään nimessä voisi hyväksyä siinä esitettyjä todistuksia. Toisaalta jos uskomme ZFC:n olevan ristiriidaton, eivät Gödelin epätäydellisyysteoreemat anna meille mitään syytä epäillä tätä ristiriidattomuutta.

Franzén (s. 105) huomauttaa myös toisesta tärkeästä asiasta: emme käytä PA:ta tai ZFC:ta pelkästään siitä syystä, että uskomme niiden olevan ristiriidattomia. Käytämme niitä, koska uskomme niiden aksioomien olevan *tosia* ja päättelysääntöjen *totuuden säilyttäviä*. Koska ajattelemme aksioomien olevan tosia, ajattelemme niiden olevan myös ristiriidattomia. Uskomme matemaattisten järjestelmien pätevyyteen on uskoa niiden aksioomien totuuteen ja päättelysääntöjen pätevyyteen. Gödelin epätäydellisyysteoreemoilla ei ole mitään vaikutusta tähän uskoon, sillä ristiriidattomaksi uskomamme systeemin sisäisellä ristiriidattomuustodistuksella ei voisi olla mitään arvoa.

*Kaikki* matematiikka olettaa jotakin. Tarkalleen ottaen jokainen matemaattinen todistus on muotoa: ”jos tietty aksiomaattinen systeemi (PA, ZFC,…) on ristiriidaton, niin…”. Mitään perimmäistä ristiriidattomuustodistusta, joka oikeuttaa kaiken matemaattisen päättelyn ei voisi koskaan olla olemassa, koska se pitäisi myös todistaa jossakin systeemissä, jonka ristiriidattomuus oletetaan. Siksi sellaisen ristiriidattomuustodistuksen *puute* ei myöskään vaikuta millään tavalla matematiikan varmuuteen – mikä on hyvin nähtävissä siitä, että matemaattisten tulosten varmuutta ei matemaatikkojen keskuudessa pidetä nykyään yhtään sen kyseenalaisempana kuin ennen Gödeliäkään.

## 7. Fysiikan epätäydellisyys

Koska empiiriset tieteet käyttävät matematiikkaa hyväksi kaikkialla, on kenties luonnollista ajatella, että matemaattisten systeemien epätäydellisyys vaikuttaa myös niissä. Erityisesti tämä koskee fysiikkaa, jonka nykytutkimuksessa keskeisenä tavoitteena on ns. *kaiken teoria*, eli yksi teoria, joka selittää kaikki fysiikan ilmiöt. Koska tällainen kaiken teoria suurella todennäköisyydellä tarvitsee aritmetiikkaa rikkaampaa matematiikkaa, se tulisi sisältämään oman Gödel-lauseensa ja olisi siten epätäydellinen. Fyysikko Freeman Dyson kirjoittaa:

> ”Toinen syy, miksi uskon tieteen olevan ehtymätön on Gödelin teoreema [josta seuraa, että] puhdas matematiikka on ehtymätön. […] Väitän, että Gödelin teoreeman vuoksi myös fysiikka on ehtymätön. Fysiikan lait ovat äärellinen joukko sääntöjä ja sisältävät matematiikan säännöt, joten Gödelin teoreema

> pätee niihin. Teoreemasta seuraa, että tietomme jopa fysiikan perusyhtälöistä tulee olemaan epätäydellistä” (Franzén 2005, s. 87)

Dyson on oikeassa siinä, että jos meillä olisi kaikenkattava ristiriidaton fysiikan teoria, joka sisältää vähintään aritmetiikan, olisi tämä teoria epätäydellinen. Mutta kuten Feferman vastasi Dysonille[3], tämä epätäydellisyys koskisi teorian *aritmeettista* osaa. Fysiikan perusyhtälöt eivät tällaisessa teoriassa voisi ratkaista kaikkia aritmetiikan lauseita, mutta ne voisivat silti olla täydellisiä mitä tulee *fysiikan* lauseisiin. Näin ollen fysiikan totuuksien joukko voisi kaiken teoriassa olla täydellinen, mikä varmasti riittäisi fyysikoille.

Entä mitä tapahtuu siinä tapauksessa, että empiirinen lause *on* oleellisesti aritmetiikan lause? Esimerkiksi lause ”miljoona litraa vetävän kuutionmuotoisen vesisäiliön särmä on kymmenen metriä” on empiirinen, mutta selvästikin se on myös aritmetiikan lause. Gödelin teoreemojen nojalla emme voi luetella kaikkia aritmetiikan totuuksia. Eivätkö tällöin myös tietyt empiiriset totuudet jää empiiristen teorioiden ulkopuolelle?

Tämä olisi totta, jos todella etsisimme *kaiken* teoriaa, siis sellaista, joka selvittäisi kaikki fysiikan ja matematiikan totuudet. Mutta fysiikassa kaiken teoriassa etsitään nimenomaan kaikkia *fysiikan* totuuksia, ja epätäydellisyysteoreemat eivät anna mitään syytä olettaa, ettemmekö voisi tietää niistä jokaista. Se, että aritmeettinen teoria kaikista tällaisista totuuksista on epätäydellinen, ei tarkoita sitä, ettei fysiikan teoria pystyisi selvittämään niitä täydellisesti. Aritmetiikka antaa tietenkin luotettavia ennustuksia empiirisistä totuuksista, mutta se, että fysiikan teoria sisältää aritmetiikan ei tee siitä epätäydellistä muuten kuin aritmetiikan osalta.

## 8. Gödeliläiset virhepäätelmät muilla aloilla

Tähän asti esitetyt epätäydellisyysteoreemojen sovellukset sijoittuvat enemmän tai vähemmän ymmärrettäville alueille, sillä niissä on järkevää tulkita keskeisiä käsitteitä formaaleina systeemeinä. Mutta Gödelin teoreemoja on tulkittu myös alueilla, jotka tuntuvat olevan kaukana niiden soveltuvuusalueista. Yksi toistuva argumentti on ollut *teologinen* tulkinta epätäydellisyydestä. Daniel Graves kirjoittaa:

[3] The New York Review of Books, July 15, 2004.

> [Gödel] näytti, että mikään systeemi ei voi ratkaista tietyn $A$:n ja ei-$A$:n välillä, edes silloin kuin tiedämme $A$:n olevan tosi. Mikään äärellinen systeemi, jolla on riittävästi voimaa sisältää täysi lukuteoria, ei voi olla itsensä ilmaisema. Judeo-kristillisyys on pitkään pitänyt totuutta pelkän järjen yläpuolisena. […] Gödeliläinen kuva sopii siihen, mitä kristityt uskovat maailmankaikkeudesta. […] seurauksena Gödelin teoreemasta, uskon näytetään olevan (lopulta) ainoa vastaus todellisuuteen. (Franzén 2005, s. 92)

Michael Guillen pukee tämän viimeisen ajatuksen sanoiksi:

> Ainoa tapa hyväksyä todistamaton totuus, matemaattinen tai muunlainen, on hyväksyä se uskonkappaleena. (Guillen 1983)

Nämä varmasti riittävät esimerkeiksi Gödelin soveltuvuudesta teologiaan. En käy tässä tarkemmin läpi näitä yksittäisiä argumentteja, jotka sisältävät monia virhetulkintoja Gödelin tuloksista. Gravesin ja Guillenin argumentti perustuu siihen, että Gödel-lauseiden totuus joudutaan uskomaan, koska sitä ei voida todistaa. Tämä ei kuitenkaan pidä paikkaansa, vaan Gödel-lauseiden totuus *seuraa* siitä, että uskomme formaalin systeemin totuuteen. Paljon parempi argumentti olisi huomauttaa, että kaikki matematiikka perustuu uskoon aksioomien totuudesta ja päättelysääntöjen pätevyydestä. Tosin tämä ei ole rinnastettavissa uskoon Jumalasta, eikä liity millään tavalla Gödelin epätäydellisyysteoreemoihin.

Tämänkaltaisten yksittäisten virheiden sijaan haluan keskittyä yleisempään ongelmaan tässä ja useassa muussa argumentissa. Puretaan Gravesin argumentti palasiksi. Sen looginen muoto on yksinkertaisesti:

Formaali matematiikka on epätäydellistä.
Inhimillinen toiminta on epätäydellistä.
On olemassa jotain täydellistä (tai: uskomme johonkin täydelliseen).
___
Tämä täydellinen on Jumala (tai: on perusteltua uskoa Jumalaan).

Kuten huomataan, formaalin matematiikan epätäydellisyys näyttelee päättelyssä hyvin yksinkertaista roolia: se toimii vain esimerkkinä siitä, että ihmiset eivät tiedä kaikkia maailman totuuksia. Periaatteessa sen siis voisi korvata millä tahansa ratkeamattoman tuntuisella lauseella, kuten kysymyksellä siitä, mitä oli ennen alkuräjähdystä. Mutta miksi käyttää epätäydellisyyttä esimerkkinä, kun tiedetään että se on suurimmalle osalle lukijoista vaikeasti avautuva ongelma?

Itse asiassa uskon motivaation olevan juuri tämä vaikeaselkoisuus: Gödelin epätäydellisyysteoreemojen käyttö saa argumentin vaikuttamaan täsmälliseltä ja syvälliseltä. Yleinen muoto näissä ja muissa teologisissa Gödeliin nojaavissa argumenteissa näyttäisi olevan seuraava: jos matematiikka on epätäydellistä, on kaikki muu vielä epätäydellisempää. Tämä itsessään on tietenkin väärin, sillä formaalia epätäydellisyyttä ei voi suoraan verrata muun inhimillisen toiminnan epätäydellisyyteen. Mutta se, että päädytään johtopäätöksiin Jumalan olemassaolosta aritmetiikan epätäydellisyyden nojalla, ei liity enää mitenkään Gödelin epätäydellisyysteoreemoihin. Tai jos liittyy, niin asia vaatisi ainakin tarkan selvityksen.

Tämä virhe on erittäin yleinen kun epätäydellisyyttä käytetään hyväksi matematiikan tai matematiikanfilosofian ulkopuolella. Parhaimmillaankin se on tapaus tutusta virhepäätelmästä, jossa epätäydellisyyden sisältöä ja soveltuvuusaluetta ei tunneta. Pahimmillaan se on suoraa älyllistä epärehellisyyttä: kirjoittaja uskoo, ettei lukija ymmärrä Gödelin epätäydellisyysteoreemoja, ja katsoo siten voivansa käyttää niitä perusteluina. En ota kantaa kummasta seuraavissa esimerkeissä on kyse, mutta tuloksena on tulkinnasta riippumatta yhtä huonoa tiedettä:

> [Gödelin tuloksesta lähtien] valtiotieteilijät ovat voineet ymmärtää miksi Lenin täytyi muumioida ja panna esille ”satunnaisten” tovereitten nähtäväksi mausoleumiin. (Debray 1980, s. 70)

Michel Serres tukee Debrayta:

> Soveltamalla Gödelin teoreemaa kysymykseen suljetusta ja avoimesta, siten kun ne suhtautuvat sosiologiaan, yhdellä eleellä Régis Debray tiivistää ja saattaa loppuun viimeisen 100 vuoden historian ja työn. (Serres 1995, s. 452)

Nyt kysymys kuuluu: mitä roolia Gödelin epätäydellisyysteoreemat näyttelevät Debrayn argumentissa? Ainoa mahdollinen vastaus on jonkinlaisena analogiana, sillä emme voi soveltaa formaalin aritmetiikan epätäydellisyyttä sosiaalisiin rakenteisiin. Mutta mitä hyötyä tällaisesta analogiasta on Debrayn ja Serresin lukijoille? Kuten Alan Sokal ja Jean Bricmont (1998) huomauttavat, tällainen päättely sotii koko analogian käytön ideaa vastaan. Tavalliselle sosiologisen tekstin lukijalle Gödelin epätäydellisyysteoreemat eivät todennäköisesti ole tuttuja. Käyttämällä niitä analogiana yhteiskunnan ilmiöistä – jotka ovat oletettavasti paljon tutumpia – Debray tuntuisi käyttävän analogioita, jotka ovat *vaikeampia* ymmärtää kuin niiden ilmaisema suora ajatus.

Sama analogian käytön vääristyminen on läsnä myös Julia Kristevan Gödel-tulkinnassa:

> Yhdistettynä siihen, mitä juuri sanottiin runollisesta kielestä, valinta-aksioomasta seuraava konstruktiivisuuden käsite selittää, miksi on mahdotonta saada aikaan ristiriita runollisen kielen avaruudessa. Tämä huomio on lähellä Gödelin huomiota epätäydellisyyden todistamisen mahdottomuudesta systeemissä siinä formalisoiduin keinoin. (Kristeva 1969, s. 189-90)

En noteeraa tässä valinta-aksioomaa koskevaa väitettä (joka on virheellinen), vaan keskityn Gödeliin. Tietenkin Gödel todisti *täydellisyyden* todistamisen mahdottomaksi, joten Kristevan yritys soveltaa epätäydellisyysteoreemaa on välittömästi tuomittu epäonnistumaan. Mutta vaikka hän olisi käyttänyt Gödeliä oikein, miten formaalien systeemien epätäydellisyys liittyy runollisen kielen avaruuteen? Ja ennen kaikkea: miten Gödelin teoreeman pelkkä mainitseminen voi auttaa kirjallisuudentutkijaa, joka oletettavasti ei ole matemaattisen logiikan asiantuntija, ymmärtämään Kristevan ajatusta?

Tarkoituksenani ei ole kritisoida yksittäisiä ajattelijoita, vaan esitellä yleisempää Gödelin väärinkäyttöä. On ikävän yleistä käyttää epätäydellisyysteoreemoja pelkästään tuomaan painoarvoa tekstille. Kaikki tässä viimeisessä osassa esitetyt katkelmat tuntuvat syyllistyvän siihen, ja ne ovat vain jäävuoren huippu. Jos Kristeva haluaa sanoa, ettei runon kielessä voida ilmaista ristiriitaa, miksi hänen tarvitsee mainita Gödel – tai valinta-aksiooma – ajatuksensa tueksi? Kyse on toki osasta laajempaa ilmiötä, jossa matematiikan ja luonnontieteen käänteentekeviä tuloksia väärinkäytetään alueilla, joihin ne eivät sovellu.[4] Mutta tuskin mitään matematiikan tulosta on väärinkäytetty enemmän kuin aritmetiikan epätäydellisyyttä.

[4] Sokal & Bricmont (1998) tarjoaa laajan kirjon tällaisia väärinkäyttöjä.

Miksi juuri Gödeliä? Uskon perimmäisen syyn tähän olevan sen voimakas epäformaali muotoilu ”on olemassa totuuksia, joita ei voida todistaa”. Tämä ajatus on helposti ymmärrettävissä, ja se vastaa monia yleisiä intuitioita inhimillisen tiedon ja toiminnan luonteista.[5] Emme osaa ilmaista kaikkea, emmekä varmasti *tiedä* kaikkea. Kaikki kiinnostava inhimillinen toiminta *on* epätäydellistä, tavalla tai toisella. Siksi on helppo ajatella olevan jokin yhteys Gödelin epätäydellisyysteoreemojen ja minkä tahansa inhimillisen toiminnan välillä. Mutta Gödelin todistama epätäydellisyys – ja siinä muotoiltu tosi lause, jota ei voida todistaa – pätee ainoastaan todistuksessa esitetyin ehdoin. Soveltaessamme epätäydellisyysteoreemoja filosofiaan näiden ehtojen on pakko täyttyä – muuten argumenttimme ovat parhaimmillaankin vain tarpeettoman hankalia analogioita.

**Kirjallisuus**


Azzouni, Jody (1999), “Comments on Shapiro”, *Journal of Philosophy*, 96, s. 541-544.

Debray, Régis (1983), *Critique of Political Reason*, englanniksi kääntänyt David Macey, New Left Books, London.

Feferman, Solomon (1991), “Reflecting on Incompleteness”, *Journal of Symbolic Logic*, 46, s. 1-49.

Field, Hartry (1999), “Deflating The Conservativeness Argument”, *Journal of Philosophy*, 96, s. 533-540.

Franzen, Torkel (2005), *Gödel’s Theorem: an Incomplete Guide to Its Use and Abuse,* A K Peters, Wellesley, Massachusetts.

Gödel, Kurt (1931), “On formally undecidable propositions”, teoksessa *Collected Works Volume I*, Oxford University Press, New York 1986, s. 145-195.

Gödel, Kurt (1951), “Some basic theorems on the foundations of mathematics and their implications”, teoksessa *Collected Works Volume III*, Oxford University Press, New York 1995, s. 304-323.


[5] Tämä vastaa David Wayne Thomasin (1995, ss. 252-253) huomiota: ”Vertaan näitä postmoderneja kriittisiä väittämiä, en Gödelin teoreemaan itseensä, vaan väittämiin, jotka voisivatkuvailla Gödelin teoreemaa”. Vaikka Thomaskin syyllistyy artikkelissaan gödeliläisiin virhepäätelmiin, edustaa hän kuitenkin varovaisempaa Gödel-tulkintaa, jossa epätäydellisyysteoreemoja käytetään puhtaasti analogioina niitä kuvailevien epäformaalien ideoiden kautta.

Halbach, Volker (2001), “How Innocent is Deflationism?”, *Synthese*, 126, s. 167-194.

Horsten, Leon (2009), “Levity”, *Mind*, 118, s. 555-581.

Horwich, Paul (1998), *Truth*, Clarendon Press, Oxford.

Isaacson, Daniel (1987), “Arithmetical Truth and Hidden Higher-Order Concepts”, teoksessa *Hart, W.D.* (ed.): *The Philosophy of Mathematics*, Oxford University Press, Ofxord 1996, s. 203-224.

Ketland, Jeffrey (1999), “Deflationism and Tarski’s Paradise”, *Mind* 108, s. 69-94.

Kristeva, Julia (1969), *Séméiotiké: Recherches pour une sémanalyse*, Éditions du Seuil, Paris.

Lucas, John (1961), “Minds, Machines and Gödel”, *Philosophy*, 36, s. 112-127.

Pantsar, Markus (2009), “*Truth, Proof and Gödelian Arguments: A Defence of Tarskian Truth in Mathematics*”, *Philosophical Studies from the University of Helsinki,* 23. Saatavilla osoitteesta: http://urn.fi/URN:ISBN:978-952-10-5374-0.

Penrose, Roger (1989), *The Emperor’s New Mind*, Oxford University Press, New York.

Penrose, Roger (1994), *Shadows of the Mind*, Oxford University Press, London.

Nagel, Ernest & Newman, James R. (1959), *Gödel’s Proof*, Routledge & Kegan Paul, New York.

Russell, Bertrand (1920), *Introduction to Mathematical Philosophy*, Second Edition, Dover, New York 1993.

Serres, Michel (1995), “Paris 1800”, teoksessa *A History of Scientific Thought: Elements of a History of Science*, käännetty ranskasta, Blackwell, Oxford.

Shapiro, Stewart (1998), “Proof and Truth: Through Thick and Thin”, *Journal of Philosophy*, 95, s. 493-521.

Smorynski, C. (1977), “The Incompleteness Theorems”, teoksessa Barwise, J. (ed.): *Handbook of Mathematical Logic*, North Holland Publishing, New York 1977, s. 821-865.

Sokal, Alan & Bricmont, Jean (1998), *Intellectual Impostures*, Profile Books, London.

Tarski, Alfred (1935), “Das Wahrheitsbegriff in formalisierten Sprachen”, *Studia Philosophica* I, s. 261-405. Englanniksi kääntänyt J.H. Woodger (1956): “The Concept of Truth in Formalized Languages”, teoksessa *Logic, Semantics, Metamathematics,* Hackett, Indianapolis 1983, s. 152-278.

Tennant, Neil (2002), “Deflationism and the Gödel Phenomena”, *Mind,* 111, s. 551-582.

Thomas, David Wayne (1995), “Gödel’s Theorem and Postmodern Theory”, *PMLA*, Vol. 110, No. 2, ss. 248-261.